

## **Reviewer's report**

**Title:**High-coverage sequencing and annotated assemblies of the budgerigar genome

**Version:**1**Date:**10 November 2013

**Reviewer:**Niranjan Nagarajan

### **Reviewer's report:**

#### Minor Essential Revisions

- 1) As this dataset can be a valuable resource for developers of assembly algorithms, plots showing the estimated fragment length distribution of the various mate-pair libraries (by mapping to the assembled genomes) would be useful to assess the quality of the various libraries.
- 2) Another assessment that would help users of the data would be to provide plots for the per base sequence quality distribution of the various libraries.
- 3) Please describe the parameter settings used for various software packages and provide more detail in places where currently a custom pipeline is being referred to.
- 4) In the section on annotation, it would be useful to clarify what is meant by "projecting" proteins onto the genome and if any de novo methods (the answer seems to be "No") were also used for the annotation.
- 5) Access to the complete optical mapping dataset (all maps rather than just aligned maps) would also add value to this unique genome sequencing dataset.

#### Discretionary Revisions

- 1) From the description in the text, the scaffold N50 seems to have been computed taking Ns into account. If this is the case, it might be more meaningful to report lengths after ignoring Ns.

**Level of interest:**An article of importance in its field

**Quality of written English:**Acceptable

**Statistical review:**No, the manuscript does not need to be seen by a statistician.

### **Declaration of competing interests:**

Other than the fact that we work on genome assembly problems I do not have any competing interests in relation to this paper.